



SPAM EMAIL CLASSIFICATION AND IDENTIFICATION

Keerththana K. ^{1,*} and Charles E.Y.A. ²

¹ Department of Physical Science, Faculty of Applied Science, University of Vavuniya, Sri Lanka.

² Department of Computer Science, Faculty of Science, University of Jaffna, Sri Lanka.

* Corresponding author email: kkeerthu236@gmail.com

Abstract: Email has become one of the most widespread ways of communication in today's society. Email spam, commonly known as junk email, spam mail, or simply spam, refers to unsolicited messages sent in large quantities through email. Even though some spam emails contain valuable information, quite often spam emails are unwanted and lead to online fraud. Hence it is necessary to filter spam emails from regular emails. An improved spam classification approach will make users' inboxes free from spam emails while not missing any potential emails. In this research work we analyzed the classification of emails into spam and legitimate emails using the contents of the email. This work further explored the classification of the spam emails based on categories such as promotion, marketing, news, security and others. This work analyzed the applicability of the word embedding approach for spam classification. Two different kaggle datasets (sms-spam-collection-dataset, spam filter) were used in this research work. This work considered a word embedding approach for text representation and multiple classifiers (LSTM, SVM). Since there are no publicly available multiclass spam classification data sets, an incremental approach is proposed to build the classifier. Both datasets were manually categorized and used to build the multiclass classifier. This work identified the Word2Vec model with SVM classifier obtained highest accuracy of 0.86, 0.87 for both datasets. As future work, this initial classifier will be used to classify the Enron spam email dataset. With a manual analysis the results will be verified and will be used to fine tune the classifier in multiple epochs.

Keywords: Machine learning, Multiclass classification, Spam classification, Spam filter, Spam identification, Word embedding

Supervised (internal) by: Ms. R. Yasotha